

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 15:18:56 : Search time 101.54 Seconds  
(without alignments) 443.824 Million cell updates/sec

Title: US-09-052-089a-1  
Perfect score: 2384  
Sequence: 1 MPRALCTICSDFFDHSRDV.....VRVKVPSLFOAKLDTFLMS 469

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	925	38.8	433	2 T30807	TRAF interacting p
2	204.5	8.6	506	2 F85016	probable RING zinc
3	187.5	7.9	263	1 S28261	centromere protein
4	183.5	7.7	1957	2 T38077	hypothetical coile
5	182.5	7.7	1290	2 A55094	chromosomal protei
6	182	7.6	425	2 T25457	hypothetical prote
7	179	7.5	1138	2 T24635	hypothetical prote
8	178.5	7.5	764	2 T51302	myosin heavy chain
9	178	7.5	477	2 UE0343	leif protein - rat
10	177.5	7.4	1133	2 T22976	hypothetical prote
11	176.5	7.4	1156	2 E69444	chromosome assembl
12	176	7.4	1156	2 B70356	myosin heavy chain
13	176	7.4	1509	1 A272324	myosin heavy chain
14	175	7.3	1690	2 T13030	microtubule bindin
15	172	7.2	1938	2 A59293	myosin heavy chain
16	172	7.2	2116	2 A26655	myosin heavy chain
17	171.5	7.2	1999	1 S21801	myosin heavy chain
18	170	7.1	532	2 S54871	M protein - Strept
19	170	7.1	978	2 A70387	conserved hypotet
20	169.5	7.1	1744	2 JH0720	tanabin - African
21	169.5	7.1	1940	1 A24822	myosin heavy chain
22	169	7.1	472	2 S43554	plasmaingogen-bind
23	169	7.1	1935	1 S06006	myosin beta heavy
24	168.5	7.1	3187	2 JC5837	364K Golgi complex
25	168	7.0	959	2 A55913	transcytosis-assoc
26	168	7.0	1164	2 T24806	hypothetical prote
27	168	7.0	1325	2 T42722	male-enhanced anti
28	167.5	7.0	1132	2 T00259	hypothetical prote
29	167	7.0	853	2 T51505	hypothetical prote

30	167	7.0	879	2 A48575	paramyosin - nemat
31	167	7.0	1201	2 A35815	myosin heavy chain
32	167	7.0	1201	2 A35815	myosin heavy chain
33	167	7.0	2411	2 B32491	myosin heavy chain
34	166.5	7.0	555	2 C96667	unknown protein, 7
35	166.5	7.0	1931	2 A59234	slow myosin heavy
36	166	7.0	388	2 A49545	plasmaingogen-bind
37	166	7.0	484	2 S46489	M1 protein precurs
38	166	7.0	1738	2 T14867	interaplin - slime
39	166	7.0	1935	2 A59286	myosin heavy chain
40	166	7.0	2649	2 A40937	bulious pemphigoid
41	165.5	6.9	1295	2 T24587	hypothetical prote
42	165.5	6.9	1790	2 S67593	transport protein
43	165.5	6.9	1961	1 A61231	myosin heavy chain
44	165.5	6.9	2139	2 T18296	myosin heavy chain
45	165.5	6.9	2954	2 T14156	kinesin-related pr

ALIGNMENTS

Query Match	Score	DB 2:	Length
Best Local Similarity	43.8%	Pred. No. 4.5e-42;	
Matches	197;	Conservative	76; Mismatches 121; Indels 56; Gaps 5;
QY	1	MPRALCTICSDFFDHSRDVAAIHGHTFHLOCLIOSFTAPSRTPCPCRIYQKRTIIN 60	
DB	1	MPRALCTICSDFFDHSRDVAAIHGHTFHHECLVRFQAPTKPCQCKRQVSTRHIIS 60	
QY	61	KLPEFLAQEEENVLDREFLKNELDNVRALQSKDKERDSQVITLRLDLEERNATVVS 120	
DB	61	RLTFVDGLDSSVGGDPESLDNEDLVRAKVNFSKREWRQKAKADDLMEVLELRKALEN 120	
QY	121	LOQALGRAEMICSTLTKOMKRYLEEQODETRKQAEAGRLRSKMKTMQIEILLQSLQLEPV 180	
DB	121	LQKQVMEKEKLSALPTQMYTVLESQHNDTPRAKEEVRRLRIKTKTFPSLDVIVLQGAQEV 180	
QY	181	EMATRMQVQSAVEQLAVYCVSLKKEYENLKEARKASGEVADKLKRDLEFSSSKLOTVY 240	
DB	181	ESMTTDMGIDQAAVEQSLICYISLKEYEDNLKGLKSSNMCEKLKREVLTSNNKLKAL 240	
QY	241	SELDQALTELSAQKQDSADKETSLSKKRLTMQLTFLNLPVAVSEVVDRLVLE--SPAP 298	
DB	241	LELTAKADDMKSLQNDLTNMEKEKLSLKKYVEFQELSLPTPTKRNEMAGLRFERCAAP 300	
QY	299	---VEVNEKLRRPSEFRDDIDLNATFDVDTPPARPSSOHGYEKLCLKSHSPIDVDP 353	
DB	301	PONSRAVSKASCLHPSPGNEIDIDLMYDVTP-----DDVR 336	
QY	354	KKICGKRKESQSLSGQSCAGEPEDELVAGAFPLFVNATILGQKQPKRPSSESSCSADVY 413	
DB	337	KRPKVPSSKRMHL-----DSVILLRLRTKYKNLVSNQK-----FV 371	

QY 414 RTGFDGLGRTKFIQPTDVMIRPLPVKPK 443  
 Db 372 QSGYDGLGRTKFIQPTDVMIRPLPVKPK 401

RESULT 2  
 F85016  
 Probable RING zinc finger protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
 C:Accession: F85016  
 R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
 Nature 402, 769-777, 1999  
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
 A:Reference number: A85001; MUID:20083488  
 A:Accession: F85016  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-506 <STO>  
 A:Cross-references: GB:NC\_001268; NID:g7267624; PIDN:CA80936.1; GSPDB:GN00140  
 C:Genetics:  
 A:Gene: AT4G01270  
 A:Map position: 4

Query Match 8.6%; Score 204.5; DB 2; Length 506;  
 Best Local Similarity 21.5%; Pred. No. 0.00082;  
 Matches 106; Conservative 88; Mismatches 189; Indels 111; Gaps 18;

QY 5 ALCTGSD---FFDHSRDVAHICGHTFHLQCLQSFEAPAS---RTCPQCRIOVGKRT 57  
 Db 10 ALCISCYEDLKPVVENLOSISA--CGHVFEHLCLQOMFECPSTNKRNCPTCKGKSLKD 67  
 QY 58 IINKLFED-----LAQEEENVLD---EFIKNELDNVRAQLSQK----- 93  
 Db 68 PC-RLIFQSSGNQDTSASDKVGIEDPVLKGEVRLBEGVONLTSALEAKKENVEV 126  
 QY 94 -----DKEKR-DSQVIIDTLPDLEENRATVVSLOQALGAEMLCSTLKKOM 139  
 Db 127 SDRKHQCNEDLKEDKVRWMLQEIISTOHHLKLKSECCIOLNOCYKLODERTVALAKEL 186  
 QY 140 KYLEQODDEFKQAOE--EAGRLNSKMKTMEOIEILLQS-----QLPEVEIMDMGVQ 191  
 Db 187 ASKLIVDLISLEDDVLAFLKLLGNNAKTKDITDVLKSLVIRNSYKELLAKCNOQLRGE 246  
 QY 192 S-AVEQLAVYCVSLKKEEYENLKEARKASGEVADLRKDLFSSRSKLTQVVSSELDOAKTEL 250  
 Db 247 ARSEKLEKALEKTEKLEKRRLELITERENRALRDINYSK---KCSITEVSEPAIES 303  
 QY 251 KSAQKDLQADKEIMSLKKRLTMLQETLNLPPVASETVDRVLVESPAVEVNLKLRPSF 310  
 Db 304 MSFPRMLSSDNK-----VEKISTPPCKLEEKDGTFTIGSC-----LNGRDSF 346  
 QY 311 RDDIDLNATDVTPPARPSSOHGYEKLCKLEKSHSPIDQVPPKTKGPKRK----- 363  
 Db 347 VSRDSDVIDVDVDDVPEPTNSGIRDMNTNIEKGDNSMVDIKINIKRDPSSVPSYNG 406  
 QY 364 -----SOLSLGQSCAGPEPELVGAPPIVNRNAILIGQKPKRPRSSSSSKYVR 414  
 Db 407 SGNIWSSSGTNRNLGRMSKHGERNE---ATP-----SLGGSVPRK-----DDLIS 448  
 QY 415 TGFDGLGRTKFIQ 428  
 Db 449 IGPDKGKGRKIKVLK 462

RESULT 3  
 S28261  
 centromere protein E - human  
 N:Alternate names: centromere 312K protein; kinesin-related protein CENP-E  
 C:Species: Homo sapiens (man)  
 C:Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 19-Jan-2001  
 C:Accession: S28261

R:Yen, T.J.; Li, G.; Schaar, B.T.; Szilak, I.; Cleveland, D.W.  
 Nature 359, 536-539, 1992  
 A:Title: CENP-E is a putative kinetochore motor that accumulates just before mitosis.  
 A:Reference number: S28261; MUID:93024922

A:Molecule type: mRNA  
 A:Accession: S28261  
 A:Residues: 1-2663 <YEN>  
 A:Cross-references: EMBL:215005; NID:g29864; PIDN:CAA78727.1; PID:g29865  
 C:Genetics:  
 A:Gene: GDB:CENPE  
 A:Cross-references: GDB:361164; OMIM:117143  
 A:Map position: 4q24-4q25  
 C:Superfamily: centromere protein E; kinesin motor domain homology  
 C:Keywords: ATP; coiled coil; microtubule binding; mitosis; nucleotide binding; P-loop  
 F:7-335/Domain: kinesin motor domain homology <KMO>  
 F:86-93/Region: nucleotide-binding motif A (P-loop)  
 F:486-2183/Domain: coiled coil #status predicted <COI>  
 F:92/Binding site: ATP (Lys) #status predicted

Query Match 7.9%; Score 187.5; DB 1; Length 2663;  
 Best Local Similarity 25.1%; Pred. No. 0.048;  
 Matches 93; Conservative 66; Mismatches 133; Indels 79; Gaps 18;

QY 55 KRTIINKLFEDLAQEEENVLDREFLKNELDNVRAQLSQKDKER-RSQVIIDTLPDRLAE 113  
 Db 1632 KMTAVN-----ETQKMGCEIHLKQEFQKLNLENITENIRLQI-----LHENLEE 1680  
 QY 114 RNATVVSLOQALGAEMLCSTLKKOMKYLEQODET---KQAOEAG---RLRSKMT 165  
 Db 1681 MR-SVTKERDRLKVE---ETLKVERDQLEKNERETITRDLERQEEELKIVHMLKEHQET 1736  
 QY 166 MEQIEELLQSLPPEVEIMDMGVQSAV-----EQLAVYCVSLKKEEYENLKEARKA 217  
 Db 1737 IDKLRGIVSEKTEFNISMQKDLNLSNDALKAQDLKIOEELRIAHMHLKEQDFTIDKRG 1796  
 QY 218 SGEVADKL---RDLFSSRSKLTQVVSSELDOAKLELKSADK---EIMSLKK- 270  
 Db 1797 VSEKTDKLSMKNQDLNLSNKKLQEKIQELKANHQITLTKLVNENPQKVSMDQLKKQ 1856  
 QY 271 -----LTMQL-ETLNLPPVASETVDRL-----VLESPAVEVNLKLRPSFDDI--- 314  
 Db 1857 KDQSLTSLKLEIENLNAQELHLENLEEMKSVMERDMLRVEETLKLERQLESLSQETK 1916  
 QY 315 --DLNATFDVTPPARPSSOH-----GYEKLCK-----LEKSHSPIDQVPPK 355  
 Db 1917 ARDEIQDQELKT--ARMLSKHEKETYDKLRKISKETIQISDIQKDLKSKDELQ---KR 1971  
 QY 356 ICKGPRRESQL 366  
 Db 1972 IQELQKKELQL 1982

RESULT 4  
 T38077  
 hypothetical coiled-coil protein - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T38077  
 R:Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 submitted to the EMBL Data Library, April 1996  
 A:Reference number: 221767  
 A:Accession: T38077  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1957 <CON>  
 A:Cross-references: EMBL:470690; PIDN:CAA94624.1; GSPDB:GN00066; SPDB:SPAC1F3.06c  
 A:Experimental source: strain 972h-; cosmid c1F3  
 C:Genetics:  
 A:Gene: SPDB:SPAC1F3.06c  
 A:Map position: 1

[illegible]

RESULT 5  
A55094  
Chromosomal protein XCAP-C - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 23-Mar-1995 #sequence\_revision 05-Apr-1995 #text\_change 02-Jun-2000  
C:Accession: A55094  
R:Hiirano, T.; Mitchison, T.J.  
Cell 79, 449-458, 1994  
A:Title: A heterodimeric coiled-coil protein required for mitotic chromosome condensation  
A:Reference number: A55094; MUID:95042742  
A:Accession: A55094  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1290 <HTR>  
A:Cross-references: GB:U13673; NID:g563811; PIDN:AAA64679.1; PID:g563812  
C:Superfamily: Chromosome segregation protein SMC1  
C:Keywords: chromosomal protein; DNA condensation; heterodimer

Query Match	7.7%	Score 182.5;	DB 2;	Length 1290;
Best Local Similarity	22.1%;	Pred. NO. 0.037;		
Matches 100; Conservative	82;	Mismatches 150;	Indels 121;	Gaps 20;

```

0Y 15 DHSRVA-AIHGHNPHLOLLOSEPAPSRCPQ-----R10VGKRIT----- 59
Db 620 DERKVALSSSCGALDH1---VVDITDA-----QECVNFLLKONGVAF15IDKKWYE 672
0Y 60 ---NKL-----FEDLAEENENVDREF-----LKNELDWVRAOLSOXDEKRD5 100
Db 673 KGLNKIQRPENIRPLEFDWVKVAKDEQIKPAFYALRDTIANNLDQATRAFCQDKRWR-- 730
0Y 101 QYIITPLRBTLEBRNATVVSLOQALGKAMLCSTLKK---QMKYIEQO----- 145
Db 731 ---VTLQOILIEQSGTMTGGGKVMKGRMSSVWEISDDLOKKENKLKTDTPATEI 787
0Y 146 ODETKOAEAGRLR-----SKKMTQIEILLQSOLEPREVEEMIRDMGVQO 191
Db 788 QORKAHLEBVALKLRQATREMKNTPEKTYASLOISEQVNH1KAOQKLELVNVAAPBK 847
0Y 192 SAVQOLAUVSVLKYEXENLKE-ARKASEV-----ADKLKDLFFSRSLQTYISL 243
Db 848 NOOKOMKMLLETLKYEYEVAAERAGVEAVERLKLIVDINNHHK1KKAQODKLDKVTKEI 907
0Y 244 DQ-----AKTEKLSAQODLOSA-----DKETSLKKKLTLMQETL----- 278

```

```

Db      908  DECA$ATITKQADV$ITTAORNLKKG$EAVART$EKEIVAND$ISELTIDLKLLEK$EATVM  967
QY      279  -----NLPVASEYTDRLVLE$P$AVE$N$NLRLR$P$FDDIDL$NATEVDVDT$PP$R$P$S  330
Db      966  NECKEAECSLPEV-GEQR$RSLLOELKAIQ$E$K$H$ALQ$EAL-NIRLWIE-QIDSHIAE-H  1022
QY      331  S$Q$HGYE$KLEKLEK$H$P$IDVPPK$KICG$P$P$KE  363
Db      1023  OSKIKYQ$KEITK$ISL$KIEIDIEP$EVL$GLA$OE  1055

```

```

RESULT      6
T25457
hypothetical protein B0432.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000
C:Accession: T25457
R:Henkhaus, J.; Wohldmann, P.
submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid B0432.
A:Reference number: Z20038
A:Accession: T25457
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-425 <HEN>
A:Cross-references: EMBL:U080836; PIDN:AA37893.1; GSPDB:GN00020; CESP:B0432.5
A:Experimental source: strain Bristol N2; clone B0432
C:Genetics:
A:Gene: CESP:B0432.9
A:Map position: 2
A:Introns: 64/3; 99/3; 165/1; 267/3; 350/2; 386/2
C:Superfamily: RING finger homology
F:188-238/Domain: RING finger homology <RRN>

```

Query Match	7.68;	Score 182;	DB 2;	Length 425;
Best Local Similarity	23.08;	Pred. No. 0.01;		
Matches 65; Conservative	55;	Mismatches 102;	Indels 60;	Gaps 10;

Qy	3	IKALITCSDFPDHSDVAALHCGHTPHLOCLISFETAPSRTPQOCXIDYQKRTIINKL	62
Db	188	LOGSCSICEFDLKQNDKISATVCGIYHHGCSQWIAI--RQPCSCRRYPKKGFEKTL	245
Qy	63	FFDL-----AOEENVDLREFLKNELNVRAOLSOQKEKRDOSVILDTLRLLEEENA	116
Db	246	FVDYQRMGSEAKPEPEIDRE-----BHYKISTSLKVEQEK-----LGTINTENKNLXD	294
Qy	117	IYVSLDQAL-----GKAEMLCSTLTKOMKYULEODDETKOAOEBAGRLSKAKKTYHEQIELL	172
Db	295	IYVSLSEKKLIRKEDYKROEIPKLOATIMHLTISSEETIYVLRLEQESNNRILKTECFYKIL	354
Qy	173	LOSQLEPEVEEMIRDQGVGSAVEAOYAVVC-----VSLKEKENLKEARKASGEVADLRLK	227
Db	355	IYH-----SSEADKOGLCEYLKKNKGNLDTEKFPQJLKMSTYK---DLTDKRE	397
Qy	228	DLFSSRSKLQIVYSELDQAKLELKSQAKDLOSADKEINSLSK	269
Db	398	-----AAKEIEQLKMEVOS-----LKRPAODEDAIRK	424

RESULT 7  
T24635  
hypothetical protein T07C4.10a - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 01-Dec-2000  
C:Accession: T24635; S41023  
R: Buck, D.  
submitted to the EMBL Data Library, February 1995  
A: Reference number: Z19915  
A: Accession: T24635  
A: Status: preliminary; translated from GB/EMBL/DBJ  
A: Molecule type: DNA



## RESULT 10

hypothetical protein F59A2.6 - *Caenorhabditis elegans*  
T22976  
C/Species: *Caenorhabditis elegans*  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C/Accession: T22976; T23157  
R/Lighting, J.  
Submitted to the EMBL Data Library, June 1994  
A/Reference number: Z19645  
A/Accession: T22976  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1133 <M1>  
A/Residues: 1-1133 <M1>  
A/Cross-references: EMBL:Z4801; PIDN:CAA84332.1; GSPDB:GN00021; CESP:F59A2.6  
A/Experimental source: clone F59A2  
R/Burton, J.  
Submitted to the EMBL Data Library, October 1995  
A/Reference number: Z19700  
A/Accession: T23157  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1133 <M1>  
A/Cross-references: EMBL:Z6514; PIDN:CAA91344.1; GSPDB:GN00021; CESP:F59A2.6  
A/Experimental source: clone K01A11  
C/Genetics:  
A/Gene: CESP:F59A2.6  
A/Map position: 3  
A/Intons: 13/2; 43/3; 107/3; 413/3; 492/3; 567/3; 635/3; 710/3; 738/3; 795/2; 1008/3; 1

Query Match 7.4%; Score 177.5; DB 2; Length 1133;  
Best local Similarity 20.5%; Pred. No. 0.059;  
Matches 102; Conservative 103; Mismatches 169; Indels 123; Gaps 22;  
QY 46 CPOCRIOVGR---RTIINKLFEDLAQEEENV-----LDREFLKNEIDNVRAQ 89  
DB 57 CDALQAEVNAKALREIEIQAKYDDVTKAEIRIGELSESKVLESEQAPENEKEOREE 116  
QY 90 LSOKDKKRS-QYIIDLTLDLTLEFNATVVS-----LQALGKAMLCSTLKKMKY 141  
DB 117 QLAAMKELKSEONILDEVTKLEQSEEVYLAANGAIQETLEKESEKSTAKTELEA 176  
QY 142 LEOQODETKAOEAGRLSKMKTM-----EOLELLIQSLPEVEEMIRMGVGO- 191  
DB 177 VSKKLDSSETSLKEFSMTIEMKIQLINCEKQKDEAVELKQ-KLEEVEKMSDVEVOKQ 235  
QY 192 -----SAVEQLAVVCSLKEKEYENLEAKKASGEVADKLKDLFSSRSKLTQVY--- 240  
DB 236 LLESTTSEMKQHAFAEIVKQ---LEFAQSS-----IENLKAENERN-LKTALDESDE 287  
QY 241 -SELDQAKLELKSQKQLOQADKEIMSLKKKLLTMLOET-----LNIPV 283  
DB 288 SSASISSETTKMEAKKLELSEKESKSELREQMRLQVHNAGQDIDLOKTMLEMAKI 347  
QY 284 ASETVDLTLESAPVAVENLKLRRSPRDIDLNATFDVDTPPARPSSQHGVEKLCLE 343  
DB 348 AKSTEDS-----KLAREQLAGELDE-NAKEDL-----KVVEE 377  
QY 344 KSHSPIQ-----DVPR-KICKGPRKRESQSL-GGQSCAGEPDEELVGFPIFVNAI 393  
DB 378 EKHGTGIORAGALDDAKKEVKVLEKEQLERQSALESSQELASQKAKIQELEKEIQNA- 436  
QY 394 LGQKQPRKPSSESSCSADVVRT---GPDGLAGRTKFIQPTDVTWIRPLPVKPKTK--VKQ 448  
DB 437 -----QKRSESELETANEMVRSIATLNSNSETEILKQKLETLDKELQARQOTEKALTE 491  
QY 449 RVKRVTVPSLFOAKLDT 465  
DB 492 EINVLTT-SLAEKEQQT 507

## RESULT 11

chromosome segregation protein (smc1) homolog - *Archaeoglobus fulgidus*  
E69444  
C/Species: *Archaeoglobus fulgidus*  
C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 02-Jun-2000  
C/Accession: E69444  
R/Klein, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod  
R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E  
Glodek, A.; Zhou, L.; Overbeek, R.; Cocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A/Authors: Overbeek, R.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes,  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch  
A/Reference number: A69250; MUID:98049343  
A/Accession: E69444  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-1156 <K1>  
A/Cross-references: GB:AE000995; GB:AE000782; MID:g2689318; PIDN:AAB89690.1; PID:g264  
C/Superfamily: chromosome segregation protein SMC1

Query Match 7.4%; Score 176.5; DB 2; Length 1156;  
Best local Similarity 21.4%; Pred. No. 0.068;  
Matches 74; Conservative 89; Mismatches 116; Indels 67; Gaps 11;  
QY 55 KRTIINKLFEDLAQEEENVLD-----REFLKNED-----NVAQLSQDKREKRD 99  
DB 671 KERNLSKIYELQREKGLFAELNRAESLKRQYDEVDRLTGMISELNRNLSIDELKRT 730  
QY 100 SQVITDLRD-----TLERNATVVSLOALGKAE-----MLCSTLK 136  
DB 731 ESGRIEELREISQSKREKENYISLKDYNKSLAMEBAGELDAEIEIFRMIRGSEVP 790  
QY 137 KOMVLEQODETKAOEAGRLSKMKTMQELLLQSQLP-----VEEMIR 185  
DB 791 KIVLELKKIKIEHQRNREILISIKTIESLEFRQELSSMQEQVYLDEIKRIDELIRR 850  
QY 186 DMVGQSAVQDLAVVCSLKEKEYENLEAKKASGEVADKLKDLFSSRSKLTQVYSELDQ 245  
DB 851 TIEGKARVEEINSELELRKEERELGKELGKRDELDLQKLRMAEERKTEAETDR 910  
QY 246 AKLELKAQKQLOQADKEIMSLKKKLLTMLOETLNLPVASETVRLVLESPAPVENVLKL 305  
DB 911 LERIKIQKERTLEAESEIAETG-----VEYPPNLPPL--EKVAK-VLD-----EVLVEL 958  
QY 306 RRPSPRDIDLNATFDVDTPPARPSSQHGVEKLCLEKSHSPIQD 351  
DB 959 ---STFGDVNLKAIQEEVYKARDELVE---KMWLEKERADILD 998

## RESULT 12

B70356  
chromosome assembly protein homolog - *Aquifex aeolicus*  
C/Species: *Aquifex aeolicus*  
C/Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 02-Jun-2000  
C/Accession: B70356  
R/Dackert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;  
V.  
Nature 392, 353-358, 1998  
A/Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.  
A/Reference number: A70300; MUID:98196666  
A/Accession: B70356  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-1156 <AOF>  
A/Cross-references: GB:AE000699; MID:g2983238; PIDN:AAC06839.1; PID:g2983243; GB:AE00  
A/Experimental source: strain VF5  
C/Genetics:  
A/Gene: xcpC  
C/Superfamily: chromosome segregation protein SMC1







A:Reference number: PN0013; MUID:91151356  
 A:Accession: PN0013  
 A:Molecule type: mRNA  
 A:Residues: 1914-1998, 'T' <SU2>  
 A:Experimental source: Brain  
 C:Superfamily: myosin heavy chain; myosin motor domain homology  
 C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotide  
 F:84-763/Domain: myosin motor domain homology <MOT>  
 F:14-181/Region: nucleotide-binding motif A (P-loop)  
 F:541-575/Region: actin binding #status predicted  
 F:653-675/Region: actin binding #status predicted  
 F:836-1276/Domain: coiled coil #status predicted <COI>  
 F:1277-1999/Region: 52  
 F:125/Modified site: N6,N6,N6-trimethyllysine (lys) #status predicted  
 F:180/Binding site: ATP (lys) #status predicted  
 F:693,703/Active site: Cys #status predicted  
 F:1916/Binding site: phosphate (Ser) (covalent) #status predicted  
 F:1943/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 7.2% Score 171.5; DB 1; Length 1999;  
 Best Local Similarity 18.7%; Pred. No. 0.24;  
 Matches 83; Conservative 65; Mismatches 140; Indels 135; Gaps 14;

QY 48 QCRIOVGKRTIINKLFFDLAEOEENVLDREF-----LKNELDNVRAQLSOK 93  
 Db 1072 ELKMQLAKKE--ELQALAAVVEEAAOKNMALKKIRELSQISELDLSEFASRNKA 1129  
 QY 94 DKEKRSQVVIIDTLRDTLEERNATVSLQALGAEMLCSTLKKOM----- 139  
 Db 1130 EKORRDGEELEALKTELELDLTSTAQOELRSKREOEVLNLTKEEAKTHEAQIQEM 1189  
 QY 140 -----KYLEQOODETKQAOEAGRLRSKMTME-----QIFLLQS-----QLPEV 180  
 Db 1190 ROKHSQAVEELAEOLDEQTKRVANLEKAKOTLENERGELANEVAVLLQGGDSEHKRRKV 1249  
 QY 181 EEMIRDMGV---GQSAVEDLAAYCVSLKKEYENLKEARKASGEVADTKRRDFFSSRSKL 236  
 Db 1250 EAQIOELQVFNNEGERVTLEADQVNTLQVLELDVNTGLLSQSDSKSKTKRPSALSQSL 1309  
 QY 237 QTVYSELDAQLELKSQKDLQASDKIMSLKKLTMLQETLNPVAVSETVDRLVLESP 296  
 Db 1310 Q-----DTQELDOENRQKLSLTSLTKQVDEEKN----- 1338  
 QY 297 APVEVNLKLRPSRDDIDLNATFEDVTPPARPSSOHGYEKLCLERS---HSPIDQV 352  
 Db 1339 -----SFRQLE-----EEEEAKNN-----LEKQIATLHAQVADM 1369  
 QY 353 PKRI-----CKGRKESQSLSGQSCAGEPDELVGAFPIFVNAILGQKQ----- 398  
 Db 1370 KKMVEDVGLCTEAEEVKRKLQKDLBGLS---QNHKEKVAAYOKLEKTKRLQOELDDL 1426  
 QY 399 -PKRPRESSCSKDVATGTGFDGL 420  
 Db 1427 VDLDHQOSACNLEKKQKFDOL 1449

RESULT 18  
 S54871  
 M protein - Streptococcus sp.  
 C:Species: Streptococcus sp.  
 C:date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 26-Aug-1999  
 C:Accession: S54871  
 R:Podbielski, A.; Melzer, B.  
 submitted to the EMBL Data Library, June 1991  
 A:Reference number: S54871  
 A:Accession: S54871  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-537 <POD>  
 A:Cross-references: EMBL:X60097; NID:g840905; PIDN:CAA42693.1; PID:g840906  
 C:Superfamily: M5 protein

Query Match 7.1% Score 170; DB 2; Length 532;  
 Best Local Similarity 21.5%; Pred. No. 0.059;  
 Matches 82; Conservative 67; Mismatches 141; Indels 92; Gaps 12;

QY 55 KRTIIN-----KLFFDLAEOE-----EENVLDREFLK-----NELDNVRAQ 89  
 Db 159 KMYVNDNSYNTQREVDLIEELGKKLKENQDLDEKTKKDEFFLGRTLRINLDDLKLG 218  
 QY 90 LS-----QKDKERDSQVVIIDTLRDTLEERNATVSLQALGAEMLCSTLKK 137  
 Db 219 LINDINIDLKHELBOEKOKAEADROT--LEAKKALEEKQISDASRSLSLRDLDASBREAK 277  
 QY 138 QMYLEEOOQETKQAOEAGR--LRSKKKTMEQIEILLQSLPVEE-----MIR 185  
 Db 278 QLENEYOKLEEEKQISDASQSLRDLDASREKQLEAYQKLEONKISEASRGGLR 337  
 QY 186 DMGVGQSAVBOLAVYCVSLKKEYENLKEARKASGEVADTKRRDFFSSRSKLQTVSELDQ 245  
 Db 338 DLDSAREAKQKQVERDLANLFAELDKYEEKQISDASRKLGRDLDSR-----E 386  
 QY 246 AKLELSAQKDLQSDAKDEIMSLKKLTMLQETLNPVAVSETVDRLVLESPAPVEYNKL 305  
 Db 387 AK---KQVEKALEANSKLALEKLNKLEES-----KLTLEKKAKELQAKLEA 432  
 QY 306 RRSFRRDIDLNA-----TFVDVTPPARPSSOHGYEKLCLERKSHSPIDQPPK 355  
 Db 433 EAKALEKELKQADELAKLRAGKASDQTPAEATPGN-----KVYPGQAPQACTKPN 485  
 QY 356 ICKGRKESQSLSGQSCAGEP 377  
 Db 486 QNKEPMKETKROLPSGTGEATNP 507

RESULT 19  
 A70387  
 conserved hypothetical protein aq\_1006 - Aquifex aeolicus  
 C:Species: Aquifex aeolicus  
 C:date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 02-Feb-2001  
 C:Accession: A70387  
 V:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;  
 Nature 392, 353-358, 1998  
 A:title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
 A:Reference number: A70300; MUID:98196666  
 A:Accession: A70387  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-978 <AOQ>  
 A:Cross-references: GB:AE000718; NID:g2983504; PIDN:AA07092.1; PID:g2983515; GB:AE00  
 C:Experimental source: strain VFS  
 C:Genetics: aq\_1006  
 A:Gene: aq\_1006  
 C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 7.1% Score 170; DB 2; Length 978;  
 Best Local Similarity 22.6%; Pred. No. 0.12;  
 Matches 79; Conservative 70; Mismatches 125; Indels 76; Gaps 13;

QY 41 APSRTCPQCRIOVGKRTIIN-----KLFFDLAEOEENVLDREF-----LKN 82  
 Db 481 SPDDTQVCGIYRGKALEVNDADGISELKNAKLEKEREIDTLTKLYAKQINSLEKE 540  
 QY 83 LDNVRAQLSOKDEKRDQVVIIDTLRDTLEERNATVSLQALGAEMLCSTLKKOMKYL 142  
 Db 541 MEKLRNEVEELRKE-----IPENLKER---IKLLEELIREKLEHKNLKYRKAL 587  
 QY 143 EEOODETKQAOEAGRLRSKMTMEQIELLSQSL--PVEEMIRDMGVGQSAVNEQALAVC 201  
 Db 588 EDKQKQKEAQAKLRAQTELETLKE-KIREKSRVLKEREKELR-----VERLEDYE 638





A:Cross-references: EMBL:232678; NID:9474769; PIDN:CAA83589.1; PID:g1333838  
C:Superfamily: M5 protein

Query Match 7.1%; Score 169; DB 2; Length 472;  
Best Local Similarity 21.9%; Pred. No. 0.058;  
Matches 82; Conservative 64; Mismatches 132; Indels 96; Gaps 12;

```

OY 65 DLAOE--EENVLDREFLK-----NELDNVRAOLS-----QDKERKDSQVIT 104
    || || || || || || || || || || || || || || || || || || || || ||
DB 118 DLNDDLEELKDKKEFYIGETLRITNELDLKGLQINDINIDLKHELEOKKAEADROT-L 176
OY 105 DTLRDTLEERNATVVSLOALGKAEMLCSPLKQMKYLEEOODETKQAOEAGR-LRSM 163
    || || || || || || || || || || || || || || || || || || || || ||
DB 177 EAEKAKLEEKQISDASRQSLRDLDAAREAKKQVLEAYQKLEBEKQISDASRQSLRDL 236
OY 164 KTMEOITELLQSOLPEVE-----MIDMGVGSANVEDLAVYCVSLKKEYENLK 212
    || || || || || || || || || || || || || || || || || || || || ||
DB 237 DASREAKKQLEAEYQKLEENKISEASRKGLRDLDAAREAKKQ-----LEAEHQKLE 289
OY 213 EARRASGEVADLRKLDLFSR-----SKIQVYSELDOAKLE----- 249
    || || || || || || || || || || || || || || || || || || || || ||
DB 290 EOKKISASAKKGLRDLDAAREAKKQVLEAYQKLEBEKQISDASRQSLRDL 349
OY 250 -----LKSADKLDQSADKEIMSKKLTMLQETLNPVASETVDRLVLESPAPVEYNLK 304
    || || || || || || || || || || || || || || || || || || || || ||
DB 350 ASREAKKQVLEAYQKLEENKISEASRKGLRDLDAAREAKKQ-----LKSLEKAELEQKLE 398
OY 305 LRRPSFDDIDLNA-----TFDVOTPPARPSSSQHGYEKLCELSHSTIQDVPK 354
    || || || || || || || || || || || || || || || || || || || || ||
DB 399 AEAKKLEKLEAKQAEELAKLRAGTASDSQTPPEATPGN-----KVPKGGQAPQAGTKP 451
OY 355 KICKPKRESQSL 368
    || || || || || || || || || || || || || || || || || || || || ||
DB 452 NQNKPKKTKRQL 465

```

## RESULT 23

S06006 myosin beta heavy chain, cardiac muscle [similarity] - rat

N:Contains: myosin ATPase (EC 3.6.1.32)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 19-Jan-2001

C:Accession: S06006; S07536; I67441; A02989

R:Kraft, R.; Bravo-Zehnder, M.; Taylor, D.A.; Leinwand, L.A.

Nucleic Acids Res. 17, 7529-7530, 1989

A>Title: Complete nucleotide sequence of full length cDNA for rat beta cardiac myosin he

A:Reference number: S06006; MUID:90016823

A:Accession: S06006

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-1935 <KRA>

A:Cross-references: EMBL:X15939; NID:956656; PIDN:CAA34065.1; PID:956657

R:McNally, E.M.; Kraft, R.; Bravo-Zehnder, M.; Taylor, D.A.; Leinwand, L.A.

J. Mol. Biol. 210, 665-671, 1989

A>Title: Full-length rat alpha and beta cardiac myosin heavy chain sequences. Comparison

A:Reference number: S07535; MUID:90133919

A:Accession: S07536

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-950, RK, 953-1935 <MCN>

R:Mahdavi, V.; Lompre, A.M.; Chambers, A.P.; Nadal-Ginard, B.

Eur. Heart J. 5, 181-191, 1984

A>Title: Cardiac myosin heavy chain isozytic transitions during development and under pa

A:Reference number: I53305; MUID:85179510

A:Accession: I67441

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1871-1935 <RES>

A:Cross-references: GB:M32698; NID:9205598; PIDN:AAA1659.1; PID:g205599

R:Mahdavi, V.; Perisany, M.; Nadal-Ginard, B.

Nature 297, 659-664, 1982

A>Title: Molecular characterization of two myosin heavy chain genes expressed in the adu

A:Reference number: A02988; MUID:82220036

A:Accession: A02989

A:Molecule type: mRNA

A:Residues: 1524-1528, 'V', 1530, 'R', 1532-1730, 'H', 1733-1783, 'K', 1785-1850, 'N', 1852-185

A:Cross-references: GB:J00752; NID:9205577; PIDN:AAA1654.1; PID:9205578

C:Superfamily: myosin heavy chain; myosin motor domain homology

C:Keywords: actin binding; ATP; cardiac muscle; coiled coil; heart; hydrolase; methyl

F:88-766/Domain: myosin motor domain homology <MOT>

F:178-185/Region: nucleotide-binding motif A (P-loop)

F:548-585/Region: actin binding #status predicted

F:655-677/Region: actin binding #status predicted

F:839-1935/Domain: coiled coil #status predicted <COI>

F:839-1279/Region: 52

F:1280-1935/Region: light meromyosin

F:129/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted

F:184/Binding site: ATP (Lys) #status predicted

F:695/705/Active site: Cys #status predicted

Query Match 7.1%; Score 169; DB 1; Length 1935;  
Best Local Similarity 22.5%; Pred. No. 0.32;  
Matches 65; Conservative 62; Mismatches 106; Indels 56; Gaps 9;

```

OY 31 LQCLIQSFETAPSRTPCQRIQYQKRTIINKLFFDLAEOEENVLDREF----- 78
    || || || || || || || || || || || || || || || || || || || || ||
DB 1143 LEIISERLEEGAGATSVQIEMNKKRAEFQKMRDL-----EATLQHEATPAALRKKHADS 1199
OY 79 ---LKNELDNVRAOLSQOKERKDSQYIIDTLADTLERNATVVSLOALGKAEMLCSPL 135
    || || || || || || || || || || || || || || || || || || || || ||
DB 1200 VAELEQIDMLQVKKQKLEKSEFKLELDVYSNMDO-----ITKAAKLEKMCRTL 1252
OY 136 KQMKYLEOODETKQAOEAGRLRSKMKP-----MEQIELLOSOLP----- 178
    || || || || || || || || || || || || || || || || || || || || ||
DB 1253 EDQNMHRSKAEFTQSVNDLTQRAKLTQENGLSRQDEKALL-SQTRCKRLTYTQO 1311
OY 179 -----EVEEMIRDMGVGSANVEDLAVYCVSLKKEYENLKAEKASGEVADLRKDLFS 232
    || || || || || || || || || || || || || || || || || || || || ||
DB 1312 LEDLKRQLKEEVAKNAALHAIQASARHDCDLLREQYEETFEAKELQRLVLSKANSEVAQM 1371
OY 233 RSKLQI-----VSELDQAKLELKSADKLDQSADKEIMSKKLTMLQET 277
    || || || || || || || || || || || || || || || || || || || || ||
DB 1372 RTKYETDAIORTLELEAKKL--AQR-LQDAEAAVAVNAKCSLEKT 1417

```

## RESULT 24

JC5837

364K Golgi complex-associated protein - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 05-Mar-1998 #sequence\_revision 13-Mar-1998 #text\_change 20-Jun-2000

C:Accession: JC5837

R:Tokli, C.; Fujiwara, T.; Sonoda, M.; Hong, H.S.; Misumi, Y.; Ikehara, Y.

Cell Struct. Funct. 22, 565-577, 1997

A>Title: Identification and characterization of rat 364-kDa Golgi-associated protein

A:Reference number: JC5837; MUID:98093490

A:Accession: JC5837

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-3187 <TOK>

A:Cross-references: DBJ:D25543; NID:9516825; PIDN:BA05026.1; PID:9516826

C:Comment: This protein plays a role in the formation and maintenance of the characte

C:Superfamily: giantin

F:49-549, 624-1176, 1238-1707, 1763-3114/Domain: coiled-coil leucine zipper #status pred

F:1165-3187/Domain: membrane anchor #status predicted <MAD>

Query Match 7.1%; Score 168.5; DB 2; Length 3187;  
Best Local Similarity 19.3%; Pred. No. 0.61;  
Matches 89; Conservative 84; Mismatches 160; Indels 127; Gaps 13;

```

OY 65 DLAOEENVLDREFLKN-----ELDNVRAOLSQOKKER-----DSQYIIDTLRDLLEER 114
    || || || || || || || || || || || || || || || || || || || || ||
DB 1310 ELSSSLKTLADLEHKTLOPELTQKHVGOKKEEVSYLVGOLGEKEQTLTVVQTEMEBO 1369

```

Qy	115	NATVYSL-----	---OALGKAMLSCTKKQKYLEOO	146	
Db	1370	ERLIKATHTOLEMOAENEHERLKOVVEITCEIKLOPKRELEESKAKOOLQRLQDALLSR	1429		
Qy	147	----DETQAOEAGRLRSKM-----	KTMEQIELLLOSOLPEVEEMIRMGVOSAVEOLA	198	
Db	1430	KEALKEKMSIQEOBLSKRDVAENHTYTSLSADVESQVSVONOEKDALLGKLLALQOEROKLI	1489		
Qy	199	V-----	---YCVSLK-----	205	
Db	1490	YEMKSLLENQSLGSGCESIKLALGTLGTEDKEKIMLEBSVBCSKIAESTEMOEKHKELQ	1549		
Qy	206	KEYENLKEARKASGEVADKIRKLPJFSRSBKLOTVYSELDOALLEYLKSQKLOQSADKIM	265		
Db	1550	KEYEVLLQSYEVNSENERRLOHVESEVROEKQVAKLSASDSKRRERKQLOADEQME	1609		
Qy	266	SLKKKL-----	TMLQETLNPVASETVORYLVESPAPEVY-----	EKLRRPSRFDID	315
Db	1610	EMKMKMKKPKAKSOKKLTLEL-----	BEENDRLRAEAOPOVGANESMALLSNASIKELLE	1665	
Qy	316	LNATFDVDPARPSSOHGYEKLCLKESHPIDQVPKKTICGPRKESQSLSGOSAG	375		
Db	1666	-RLTEKYLTSKE-----	-FALMAEKNTLSEETRNKLQVQOELKQASLTETKESD	1716	
Qy	376	EPDELYGAFPIFYRNAILGOKOPKPRBSGSKUVVAT	415		
Db	1717	EPKVDIEE-----	-VTEAVYKSOEOOSLSBNNKLEDEAENT	1751	

RESULT 25  
A:55913  
transcytosis-associated protein p115 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 23-Mar-1995 #sequence\_revision 05-Apr-1995 #text\_change 05-Nov-1999  
C:Accession: A55913  
R:Barroso, M.; Nelson, D.S.; Szul, E.  
Proc. Natl. Acad. Sci. U.S.A. 92, 527-531, 1995  
A:Title: Transcytosis-associated protein (TAP)/p115 is a general fusion factor required  
A:Reference number: A55913, MUID:95132633  
A:Accession: A55913  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-959 <BAR>  
A:Cross-references: GB:U15589, NID:9558474, PIDN:AAC52151.1, PID:9558475  
:Keywords: membrane fusion; membrane trafficking

Query Match	7.0%;	Score 168;	DB 2;	Length 959;
Best Local Similarity	20.2%;	Pred. No. 0.15;		
Matches	76;	Conservative	56;	Mismatches 93;
				Indels 152;
				Gaps 11;

Qy	52	OYGRKTIINKLFE-----DLAQEENYIDREFELK--NELDNV--RAQLS	91
Dp	571	RIGKENITLEKJGFSKHNELYPASCKPODPNPPSEMTIPDHFETTKLVKELEGVITKAIYK	630
Qy	92	QKDKERDSQVLIIDRLPTLEERNATVVSLOALCAEMICSTLKKOMKYLEQOODE---	148
Dp	631	SSEEDKKKEEV-----KTLLEGHDNIYVHKVNIYEQDLLEELKQVSYTLKCAQNEQLOT	685
Qy	149	--TKQAQ-----EAGRLSKRMKMTQEI	169
Dp	686	AVTQOASOIQOHKDOYNLLKVLQKDNHHQGSHPDQAVNGIQPEIISRLREIEELRSH	745
Qy	170	ELLQSSOLPEVBEHMRDM-----GVGQSAV-----EQLAIV-----	200
Dp	746	OYLLQSLAEKPTVLEIENLRSSQVSNSEMQALATCSPRDAEQVAELKQELSAKSOQLCSQS	805
Qy	201	-----CVSLKKEENYLNKEARK	216
Dp	806	LEITRLQTEENRELQORAEFLAKSVPEGSESELVTAKTTDVEGRSLALLQETRELKNEIK	865
Qy	217	ASGEVADKLKRDJFSRSRSLQTVYSELDQAKLELKSACQD-----IQSADKEIWSLKK	269

Db 866 ALSEERTSIOQLDSSNSTIAIIQTEKDKLYLEVYDSKKEDDILVLLADQDQIISLKS 925

QY 270 KLTMQETLNLPIVASE 286

|| | || |

Db 926 KKLQDLGH-----PVEEE 937

```

RESULT      26
T24806
hypothetical protein T10G3.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24806
R:Burton, J.
submitted to the EMBL data library, October 1996
A:Reference number: Z19937
A:Accession: T24806
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1164 <MIL>
A:Cross-references: EMBL:Z81118; PTDN:CA03330.1; GSPDB:GN00023; CESP:T10G3.5.
A:Experimental source: clone T10G3
C:Genetics:
A:Gene: CESP:T10G3.5
A:Map position: 5
A:Introns: 15/3; 73/3; 387/3; 412/3; 708/3; 818/3; 894/3; 982/3; 1080/2

```

Query Match	7.0%	Score 168;	DB 2;	Length 1164;
Best Local Similarity	24.3%	Pred. No. 0.19;		
Matches	62;	Conservative	66;	Mismatches 109; Indels 18; Gaps 8;

  

Oy	31	LOCIQSEIAPASRCPQPCROIOVGKRTIINKLFTDLQAEENVLDREFIK-NELDNVRAQ	89
		: : : : :   : : : :   : : : :   : : : :   : : : :	
Db	667	LEASVSSRDITVSTIESESELKGLTFESNPSRIETELKQVEKVSNDISEKQOEVEVLMAE	726
Oy	90	LSQKV--KERRD--SOVIITDLTDEERNAVAVSLQALGKAMLCSTLKKQKKYAE	143
		: : : :   : : : :   : : : :   : : : :   : : : :   : : : :	
Db	727	MMDKRAHKKTRKDEEAOMLRN-QEDNEFASSTLKSVDQLMKKETGEGEKNQLISK	784
Oy	144	QOODETKQAQAEAGRLSRKMKTMEOIELLQSOQLPEVEEMIRDMGVGQSAVEQOLAVYCS	203
		: : : :   : : : :   : : : :   : : : :   : : : :   : : : :	
Db	785	SQLELTKTEVERLT-IRSEEEKTOEIEIKLSKVATTAQOE-RDELTAATS-ESLRTTECN	838
Oy	204	LKKEEENLEKAAKGAQVADLRKQDLFSRSKTLQVYSSVLDQAKLELKAQDLQADGAE	263
		: : : :   : : : :   : : : :   : : : :   : : : :   : : : :	
Db	839	LNSKQSIQIESRRAHAEKQSENVLEKMTPEKSALE---KRIEERESTIQSIQALELTKDNE	895
Oy	264	IMSLSKKLTMLOETL	278
		: : : :   : : : :   : : : :   : : : :   : : : :   : : : :	
Db	896	IESLTKTORVAVDEL	910

RESULT 27  
T42722 male-enhanced antigen-2 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000  
C:Accession: T42722  
R:Kondo, M.; Saitou, S.  
DNA Seq. 7, 71-82, 1997  
A:Title: Cloning and molecular characterization of cDNA encoding a mouse male-enhanced  
A:Reference number: Z2242; MUID:97217683  
A:Accession: T42722  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1325 <KON>  
A:Cross-references: EMBL: D78270; NID: d1096175; PID: d1020389; PIDN: BAA19612.1  
A:Experimental source: Strain CD-1  
C:Function:  
A:Description: supposed to play some role for spermatogenesis  
C:Keywords: leucine zipper







```
QY      290 RLVESPAPEVNNKLRPRSPDDIDLNATEDVDPPARPSSQHGYYEKCLEKLSHP1 349
          :|         |         |         |         |         |         |         |         |         |
DB      431 STLLRFOSLEKEKHKSTERYT-VEATKKLLT-----AKRCVEMKREK 476
          :|         |         |         |         |         |         |         |         |         |
QY      350 QDVPRICKGPKRESQLSGSCAGEPDEELVGAFPIFVENAILGOKRPRSSESC 408
          :|         |         |         |         |         |         |         |         |         |
DB      477 DEMAKE-----KEEVKRLLEGQVREEKEKEK-----LMEITLGLGEERREALROLC 523
          :|         |         |         |         |         |         |         |         |         |

RESULT   35
A59234    slow myosin heavy chain 3 - quail
C:Species: Coturnix coturnix
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 08-Sep-2000
C:Accession: A59234
J:Nikovits Jr., W.; Wang, G.F.; Feldman, J.L.; Miller, J.B.; Wade, R.; Nelson, L.; Stock
J Biol. Chem. 271, 1704/-17056, 1996
A>Title: Isolation and characterization of an avian slow myosin heavy chain gene express
A:Reference number: A59234; MUID:96291845
A:Accession: A59234
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1931 <NT>
A:Cross-references: GB:U33862; NID:g1289513; PIDN:AAC5912.1; PID:g1289514
Superfamily: myosin heavy chain; myosin motor domain homology
E:81-761/Domain: myosin motor domain homology <MMO>
```

Query Match	7.0%;	Score 166.5;	DB 2;	Length 1931;
Best Local Similarity	21.0%;	Pred. No. 0.43;		
Matches 75; Conservative			Mismatches 131;	Indels 77; Gaps 11

```

0Y 1A0EENYNLDREFLKNELDNVAQISQ--K0KEKRD50VI-----ID1RLDTLEE- 113
Db 1258 IEE50RTVJDLST0RAK10TENSELSR0EEKEAFIN1MRGKLTY00LEDKRL0EEB 1317
0Y 114 ---RNATVYSL0A0AGKAMELCS1LKMK0KYL0EOODEFK0Q0EAGRLR5KMT- 165
Db 1318 AKAKNALAHAL0SA0HDCCLLE0Y0EEBEMAKAE10RALSKN5EVA0MRTYKEDTAIOR 1377
0Y 166 ---ME0IELL0S0LPEYBEMIRDMGVG0SAVE0LAVYCVSLK-----EYENLK--- 212
Db 1378 TEELEAKKKLAL0RI0EAEE-----AYEAVNAKCSSLEKTKHKL0NE1ED1MAAY 1427
0Y 213 EARKASGEVADRLKRLD5SSRSK10TVYSELD0AK1ELK5AOKD10SADKE1MSLKKL- 271
Db 1428 ERSNMAAALADK0KRNFDK1ISEWK---QKPE50Q1EL5A0SK5EARS1TELEFKLNAYE 1484
0Y 272 -----1M10ETLNPVPA5ETVDR1LYESAPV9VIN1KLRRP5RDD1D1LNATFVDVT 324
Db 1485 ESLEHLEFFKRENKN10E0E15DLT0LGAS0KS1HELE-KVRK0LDAEKL1E0A----- 1537
0Y 325 PPAR55SQHGYYEKLCELEKSHSP10DVPK1CK0GRKKS0LS1G0SQSAGPDEFL 381
Db 1538 -----ALEEAPASLEHEGK1TLR0LE1N0YKAVEYERK1AEKDEEM 1578

```

RESULT 36  
A49545  
plasminogen-binding protein PAM precursor - Streptococcus pyogenes (fragment)  
N:Alternate names: plasminogen-binding M-like protein (Pd 53)  
C:Species: Streptococcus pyogenes  
C:Date: 23-Mar-1995 #sequence\_revision 05-Apr-1995 #text\_change 26-Aug-1999  
C:Accession: A49545, S61084, S60829, S70459, S32619  
R:Berger, A.; Sjoerding, U.  
J. Biol. Chem. 268, 25417-25424, 1993  
A:Title: PAM, a novel plasminogen-binding protein from Streptococcus pyogenes  
A:Reference number: A49545; MUID:94064605  
A:Accession: A49545  
A:Molecule type: DNA  
A:Residues: 1-388 <BER>  
A:Cross-references: EMBL:Z22219; NID:g288978; PIDN:CAA80222.1; PID:g940870  
R:Whitmore, A.; Kapur, V.; Sullivan, D.; Musser, J.; Kehoe, M.

submitted to the EMBL Data Library, July 1994

A:Description: Noncongruent relationships between variation in emm1 gene sequences and

A:Reference number: S61072

A:Accession: S61084

A:Molecule type: DNA

A:Residues: 13-96 <EMHA>

A:Cross-references: EMBL:U01975; NID:9533627; PIDN:AAA99591.1; PID:91235829

A:Experimental source: serotype M53

R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.

Mol. Microbiol. 14, 619-631, 1994

A:Title: Non-congruent relationships between variation in emm gene sequences and the

A:Reference number: S60784; MUID:95198537

A:Accession: S60829

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 17-77 <EMW>

A:Cross-references: EMBL:U01975

A:Experimental source: serotype M53

R:Carlsson Wistedt, A.; Ringdahl, U.; Mueller-Esterl, W.; Sjoebring, U.

Mol. Microbiol. 18, 569-578, 1995

A:Title: Identification of a plasminogen-binding motif in PAM, a bacterial surface pr

A:Reference number: S70457; MUID:96342385

A:Accession: S70459

A:Molecule type: DNA

A:Residues: 30-162 <CAR>

C:Superfamily: M5 protein

C:Keywords: cell wall

F:1-29/Domains: signal sequence (fragment) #status predicted <SIG>

F:30-388/Product: plasminogen-binding protein PAM #status predicted <MOT>

Query Match	7.0%	Score 166;	DB 2;	Length 388;
Best Local Similarity	20.2%	Pred. NO. 0.066;		
Matches 73; Conservative	62;	Mismatches 120;	Indels 106;	Gaps 11;

[illegible]

RESULT 37  
546489  
M1 protein precursor - Streptococcus pyogenes  
C:Species: Streptococcus pyogenes  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 07-May-1999  
C:Accession: S46489; S46490  
R:Kjesson, P.; Schmidt, K.H.; Cooney, J.; Björck, L.  
Biochem. J. 300, 877-886, 1994  
Article: M1 protein and protein H: IgGFC- and albumin-binding streptococcal surface p



A:Reference number: 546489; MUID:94280417  
A:Accession: 546489  
A:Molecule type: DNA  
A:Residues: 1-484 <AK>  
A:Experimental source: strain 40/58, serotype M1  
A:Accession: 546490  
A:Molecule type: protein  
A:Residues: 42-51 <AK>  
A:Experimental source: strain 40/58, serotype M1  
C:Genetics:  
A:Gene: emm1  
C:Superfamily: M5 protein  
C:Keywords: transmembrane protein  
F:1-41/Domain: signal sequence #status predicted <SIG>  
F:42-484/Product: M1 protein #status experimental <MAT>  
F:459-477/Domain: transmembrane #status predicted <TM>

Query Match 7.0%; Score 166; DB 2; Length 484;  
Best Local Similarity 21.6%; Pred. No. 0.086;  
Matches 79; Conservative 74; Mismatches 159; Indels 54; Gaps 12;

QY 64 FDLAQEEENVLDRFLKNELDNVAQLSOK-DKEKRSQVITIDRLDTLEERNATVVSLO 122  
D 119 YDLAKESTS-WDRQLEKEEKEKEALELAIDQASRDYHRATALEKELEKKKALELAID 177  
QY 123 QALGKAMLCSTLKKKKYLEQOODETK---QAOEFGRLRSKM--TMOIQLLLOSO 176  
D 178 QA-SQDYNRAVLEKELETITREOEINNNLGNKLELDQSSKEQDTTEKALEEKO 236  
QY 177 LPEV--EMIRDMGVGSAVEQLAVCVSLKKEKENLKEARKASGEVADKLKDLFSSRS 234  
D 237 ISDASRQSLRRDLDASBEAKQVEKDLANLALDKVKEKQIDASRQGLRDLASRE 296  
QY 235 KLQTVYSELDOAKLELKSADKDLQSADEITMSL-----KKKLTMLQETLNLPVA 284  
D 297 AKKQVEKDLNLTAELDKVEKEQISDASRQGLRDLASBEAKQVEKALEEANSKLAA 356  
QY 285 SEVDRLVLES-----PAPEVNLKLRPSFRDIDLNL-----TFDVPDPPA 327  
D 357 LEKANKLESKLTTEKEKALELQAKLEAKALEQOLAKQAEELAKLRAGASASQTPDT 416  
QY 328 RPSSSQHGYYEKLCLKESHSPIQDVPKIKCGPRKESQLSLGSGSCAGEPDEELVGAFPI 387  
D 417 KPGN-----KAVPGKGQAPOAGTKPNQNKAPKTKRQL--PSNGET-----ANPF 460  
QY 388 FVRNAI 393  
D 461 FTAAL 466

## RESULT 38

T14867  
Interaptin - slime mold (Dictyostelium discoideum)  
C:Species: Dictyostelium discoideum  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 11-May-2000  
C:Accession: T14867  
R:Rivero, F.J.; Kusner, A.; Brokamp, R.; Matzner, M.; Noegel, A.A.  
J. Cell Biol. 142, 735-750, 1998  
A:Title: Interaptin, an actin-binding protein of the alpha-actinin superfamily in Dictyos-  
ts.  
A:Reference number: 218248; MUID:98365468  
A:Accession: T14867  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-1738 <RIV>  
A:Cross-references: EMBL:AF057019; NID:93549260; PID:93549261; PIDN:AAC34582.1  
C:Genetics:  
A:Gene: abpd  
A:Introns: 173/2; 1680/1

Query Match 7.0%; Score 166; DB 2; Length 1738;

Best Local Similarity 20.7%; Pred. No. 0.4;  
Matches 82; Conservative 81; Mismatches 149; Indels 84; Gaps 14;

QY 75 DREFLKNELDNVRAQLSOKDREKDSQVITIDRLDTLEERNATVVSLOALGKEMLCST 134  
D 1296 DLNQLKQENKEKQLESEKKEKLOSLQDNLQNDQDKKNEKLEKEEOL----- 1346  
QY 135 LKQMKVLEQOODETKQAOEAGRLRSKMKTMO-----IELLOSQPEVEMIRDMGV 190  
D 1347 LKLOQDRNDQSSQLKQLEELKSEKNOQOLKQENELNQLNQOQSSNEITQOLKDLK 1406  
QY 191 QSAVEQLAVCVSLKKEENLKEARKASGEVADKLK-----DLFSSRSKLQTVSEL 243  
D 1407 QOQOEEQ-----QENNNKEIEIRLQEI-EQLQOQOELDQSELNKEIKIQTQOEF 1456  
QY 244 DQA-----KLEKSAQKDL---QSADEITMSLKKLT---MLQETLNLPVASETV 289  
D 1457 DLSHNSRKDQLHQLQOELDQLKQSFDDQHQFVIDERYNLQQLDQSTLSNQDL 1516  
QY 290 RLVESEPAPVEVNLKLRPSFRD-----DIDLNATFDVDPDPAPRSS---SQHGYYEK 339  
D 1517 QLEKSLKPLELDSNEKQKTTIDLLSNISNQLSLODKDLISRNNSIKTLESRTQOL 1576  
QY 340 LCLEKSHSPIQDVPKIKCGPRKESQLSLGSGSCAGEPDEELVGAFPIFVRNALILQKOP 399  
D 1577 SLDEKKNLILKDLQ-----QKQOQOQPPYASSPSSPSLSSTP-----TPKP 1621  
QY 400 KRPPSES-----SCSKDVNRTGDLGRTKF 426  
D 1622 QRPNOIEIDRLVNEIVNRNDLIRK-----NKTKF 1651

## RESULT 39

A59286  
myosin heavy chain beta chain, cardiac - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 08-Sep-2000  
C:Accession: A59286  
R:Ko, Y.L.  
submitted to Genbank, October 1996  
A:Reference number: A59286  
A:Accession: A59286  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-1935 <KO1>  
A:Cross-references: GB:075316; NID:q1698894; PIDN:AAB37320.1; PID:91698895  
A:Experimental source: strain domestica  
C:Superfamily: myosin heavy chain; myosin motor domain homology  
C:Keywords: cardiac muscle; heart  
F:88-766/Domain: myosin motor domain homology <MMO>

Query Match 7.0%; Score 166; DB 2; Length 1935;  
Best Local Similarity 22.2%; Pred. No. 0.45;  
Matches 74; Conservative 62; Mismatches 131; Indels 66; Gaps 9;

QY 79 LKNELDNVAQLSOKDREKDSQVITIDRLDTLE-----ENATVVSLOALGKRAEMLCST 134  
D 1287 LSRQLEKEALISQLTGKLTYYQQLDLKRLQEEVKAKNALNALHQLSARHADLRLQ 1346  
QY 135 LKQMKVLEQOODETKQAOEAGRLRSKMKT-----MEQIELLOSQPEVEEMIR 186  
D 1347 YEETETKALEQLRLKANSFVAMQRTKTYETDAIQRTELEAKKKLAQLQDAE----- 1402  
QY 187 MGVGQSAVEQLAVCVSLK-----EYENLK---EARKASGEVADKLKDLFSSRSKL 236  
D 1403 -----AVEAVNNAKCSLSEKTKHRLQNEIEDLMDVRSNAALALDKKORNEFDKILAEW 1456  
QY 237 QTVYSELDOAKLELKSADKDLQSADEITMSLKKLT-----TMOETLNLPVASETV 288  
D 1457 KQREK---SQSELESSQKARSLSLTELFLKNAVEESLEHLETSKKNENLDEEISDLT 1513  
QY 289 DRVLSEPAPEVNLKLRPSFRDIDLNATFDVDPDPAPRSSQHGYYEKLCLKESHP 348

```

Db      1514 EQLSSSGKTHLEL-KVRKOLEAEKLELOS-----ALESEAS 1550
QY      349 IODVKKICKPKRESQLSGGSCAGEPDEEL 381
Db      1551 LEHEEGKILNALEFNOIKAEKLEAEKDEEM 1583

RESULT 40
A:Accession: A40937
bullous pemphigoid antigen 1 precursor - human
N:Alternate names: 230k bullous pemphigoid autoantigen; BP230 antigen
C:Species: Homo sapiens (man)
C:Date: 17-Jul-1992 #sequence, revision 11-Jul-1997 #text, change 21-Jul-2000
C:Accession: I56317; A40937; A40882; A61393; A37342; S46669
R:Flight: G.W., Stanley, J.R.
J: Invest. Dermatol. 101, 244-246, 1993
A:Title: Cloning of the 5' cDNA for the 230-kD bullous pemphigoid antigen by rapid amplification of cDNA ends
A:Reference number: I56317; M0ID:93346806
A:Accession: I56317
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2649 <RES>
A:Cross-references: GB:LI1690; NID:9402479; PIDN:AAA52288.1; PID:9403124
R:Sawamura, D.; Li, K.; Chu, M.L.; Uitto, J.
J: Biol. Chem. 266, 17784-17790, 1991
A:Title: Human bullous pemphigoid antigen (BPAG1). Amino acid sequences deduced from cDNA
A:Reference number: A40937; M0ID:92011453
A:Accession: A40937
A:Molecule type: mRNA
A:Residues: 1-850, 'G', 852-1643, 'T', 1645-2363, 'T', 2365-2494, 'V', 2496-2542, 'K', 2544-2649 <
A:Cross-references: GB:M69225
R:Ranka, T.; Parry, D.A.D.; Klaus-Kovtun, V.; Steinert, P.M.; Stanley, J.R.
J: Biol. Chem. 266, 12555-12559, 1991
A:Title: Comparison of molecularly cloned bullous pemphigoid antigen to desmoplakin I cDNA
A:Reference number: A40882; M0ID:91286285
A:Accession: A40882
A:Molecule type: mRNA
A:Residues: 555-2649 <TRAN>
A:Cross-references: GB:M63618; NID:9179518; PIDN:AAA5606.1; PID:9179519
R:Sawamura, D.; Li, K.H.; Nomura, K.; Sugita, Y.; Christiano, A.M.; Uitto, J.
J. Invest. Dermatol. 96, 908-915, 1991
A:Title: Bullous pemphigoid antigen: cDNA cloning, cellular expression, and evidence for
A:Reference number: A61393; M0ID:91258872
A:Accession: A61393
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1722-2649 <SAW2>
A:Note: the sequence from Fig. 3 is inconsistent with the nucleotide sequence from Fig.
R:Owaribe, K.; Kartenbeck, J.; Stump, S.; Magin, T.M.; Krieg, T.; Diaz, L.A.; Franke, W.
Differentiation 45, 207-220, 1990
A:Title: The hemidesmosomal plaque. Characterization of a major constituent protein as a
A:Reference number: A37342; M0ID:91216368
A:Accession: A37342
A:Molecule type: mRNA
A:Residues: 1595-1942, 'R', 1944-2202 <OMA>
A:Cross-references: GB:X58677; NID:936094; PIDN:CAA41528.1; PID:936095
R:Hopkinson, S.B.; Jones, J.C.R.
Biochem. J. 300, 851-857, 1994
A:Title: Identification of a second protein product of the gene encoding a human epidermal
A:Reference number: S46669; M0ID:94280413
A:Accession: S46669
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1897-2081 <HOP>
A:Cross-references: EMBL:U04850; NID:9451556; PID:9451557
C:Genetics:
A:Gene: GDB:BPAG1
A:Cross-references: GDB:125207; OMIM:113810
A:Map position: 6p12-6p11
A:Introns: 2016/3
A:Note: the list of introns is not complete
C:Keywords: basement membrane; extracellular matrix; glycoprotein; membrane protein; pol

```

```

F:1-43/Domain: signal sequence #status predicted <SIG>
Query Match          7.0%: Score 166; DB 2; Length 2649;
Best Local Similarity 20.0%: Pred. No. 0.66;
Matches 97; Conservative 88; Mismatches 181; Indels 120; Gaps 19;

QY      33 CLIQSFETAPSPRCPCRIQGVKRTIINKLFPLDAQEEENVLDREFLKNELNDVNR----- 88
Db      1330 CRENALPVCPITQATSCRAVVG-----LQGEHDKQAEELKQAVDELTPANKR 1377
QY      89 -----QLSQDKERKDSQVIIDTLRLDLEERNATVVSLOQALGKAEMLCSTLKRO 138
Db      1378 AEQDMRELTFELNALOLEKTSSEKARLLKDKLDENNTNLRCLKLELRKDAQEKYSQO 1437
QY      139 MKYLEQOQDET-----KQAGEACRLSKKMTMQIEL-----LLQSLPVEEMIRD 186
Db      1438 LRELGRQLNOTGTGKAEMOEAASDLK-KIKRNYOLELESLNHEKRLQR---EVDRIITRA 1493
QY      187 MGVGQSAVEOL--AVYCVSLKREYENLKEARKKASGEVADLRKDLFSSRSKLTQVYSELD 244
Db      1494 HAAEKNIOHLNSQIHSFRDEKLERLQICRKSDDLKEQFER---SHEQLLONITAEKE 1550
QY      245 QAKLELKSAGKDLQADKEIMSLKKLTMLQETLNLPPVASETV---DLVLESAPAYE 300
Db      1551 NND-KIQRLNEBELKESMECAEMLKQVEELTRQNNETKLMQRLQAESENIYVLEKQTIQ 1609
QY      301 --VNLKLRPSFDDIDLNATFPVDTPPARPSSQHGTYEKLCLLEKSHSPITQVPPKICK 358
Db      1610 RCBALKTQADGDFKDL-----RSTNEH-----LHKRTQEDQDFORKI-- 1646
QY      359 GPKRESQLSGGSCAGE---PDELVGAPP-----IFRNAIIGOKQPKRPSRSSC 408
Db      1647 -----KLEEDLAKSONLVSEFKQCDQONITITQNT---KKEVRNLAEELNA 1690
QY      409 SKDVVRTPGFDGLGRTKFTQPTDTVMIRPLPVKPKTKQKQRVVKTVP-----LFG 460
Db      1691 SKEEKRRGEQKVLQQAQVQVDELNNRL-----KKVQDELHLKTEDEQMTKRKNVLFQ 1741
QY      461 AKLDTF 466
Db      1742 EESGKF 1747

```

Search completed: September 4, 2002, 16:12:45  
 Job time: 3229 sec

3

---